

**INSTITUT DE RECHERCHE
EN INFORMATIQUE ET EN AUTOMATIQUE**

**DATA ANALYSIS,
LEARNING SYMBOLIC AND
NUMERIC KNOWLEDGE**

*(Proceedings of the Conference on Data Analysis,
Learning Symbolic and Numeric Knowledge)*

**Antibes
September 11 - 14, 1989**

Edited by E. Diday

IRIA

Nova Science Publishers, Inc.

USING THE BOOTSTRAP AND THE RV COEFFICIENT IN THE MULTIVARIATE CONTEXT

Susan Holmes
Unité de Biométrie
ENSA.M-INRA-USTL
9, Place Pierre Viala
34060 Montpellier (France)

Key words : Bootstrap, Pivot, Confidence Interval, Confidence Region, Principal Components, RV-coefficient

Abstract : When bootstrapping in the multivariate context several difficulties arise, one of these is the need to find an appropriate vector space where the different bootstraps can be studied. One solution to this problem is provided through the analysis of a cube of data as in Conjoint Analysis. This method uses the concepts of characteristic operators and their inner product which will be defined in the first part of this paper. Another possibility is to bootstrap a uni-dimensional statistic, namely the RV coefficient, this provides a way of choosing the number of components to be retained in PCA, without the indecision of graphical procedures and the distributional assumptions in the usual parametric procedures.

Résumé : Cet article présente quelques développements autour de l'application du principe du bootstrap (ou méthode à la Cyrano) dans le cadre multi-dimensionnel en soulignant leur application à la construction de régions de confiance, dans les cas paramétrique simple et non-paramétrique.

L'Analyse en Composantes Principales est prise comme exemple type de procédure à laquelle peut s'appliquer la construction d'intervalles de confiance pour les valeurs propres ainsi que de cônes de confiance pour les vecteurs propres, enfin une application inédite au coefficient de corrélation RV montre l'utilité de la méthode dans le cas de statistiques dont on ignore la distribution asymptotique.

AMS : 62E25,62G15,62H20,62H25.

*visiting Stanford University for the Spring Quarter

1 Bootstrap: Motivations and Notations

Given a probability distribution F on \mathbb{R}^k , using a sample of random variables drawn from F $\mathcal{X}_n = (X_1, \dots, X_n)$ allows the construction of the empirical cumulative distribution function defined as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n 1_{(-\infty, x]}(X_i).$$

Glivenko-Cantelli's theorems ensures us that F_n is a consistent estimate of F in the following sense:

$$\sup_x |F_n(x) - F(x)| \xrightarrow{p.s.} 0 \text{ quand } n \rightarrow \infty$$

The idea behind the bootstrap is that if one wants to know the cumulative distribution function $J_n(\cdot, F)$ of a random variable $T_n(\mathcal{X}_n, F)$ depending both on the sample and on its distribution, then one may consider the approximation provided by the empirical cumulative distribution function $J_n(\cdot, F_n)$ of $T_n(\mathcal{X}_n^*, F_n)$ where \mathcal{X}_n^* is a random sample of observations which are independent and identically distributed according to F_n . Using Monte Carlo methods one can draw with replacement n observations from \mathcal{X}_n , thus constructing B such resamples provides B values of $T_n(\mathcal{X}_n^*, F_n)$, thus providing a way of estimating the cumulative distribution function $J_n(\cdot, F_n)$ only containing a simulation error. This function is then used as an estimate for the cumulative distribution function of $T_n(\mathcal{X}_n, F)$.

The theoretical justification of such a procedure has to be provided in each particular case and usually depends on the Triangular Array Convergence Property cited below, and the verification of the following **Theorem** (BERAN 1984 [1]).

Triangular Array Convergence : The class \mathcal{F} obeys the triangular array convergence property for the distance d defined on \mathcal{F} if for every sequence of c.d.f.'s $\{F_n, F_n \in \mathcal{F}\}$ such that $\lim_{n \rightarrow \infty} d(F_n, F) = 0$, we have the weak convergence of $J_n(\cdot, F_n)$ towards $J(\cdot, F)$ (which must not depend on F).

Theorem : Under the condition that the family \mathcal{F} , obeys the **Triangular Array Convergence Property** for a distance d and that the sequence $\{\hat{F}_n\}$ verifies $d(\hat{F}_n, F) \xrightarrow{p} 0$, for $\hat{F}_n \in \mathcal{F}$ then

$$\sup_x |J_n(x, \hat{F}_n) - J_n(x, F)| \xrightarrow{p} 0$$

The advantage of this resampling procedure is that it allows construction of confidence regions even in a distribution-free context the first articles published on the subject are due to EFRON 1979 [10], 1981 [11].

A more theoretical justification of the procedure as developed in sections 2 and 4 can find their source in work done by P. HALL [15], 1987 [16], and R. BERAN 1987 [2]. Such refinements as smoothing, usage of the Bayesian paradigm, importance sampling have also

enhanced the method in recent years. (Some references may be found at the end of this presentation).

We are going to present an example of how such a procedure can be used in Principal Component Analysis. However in this case the application is not a straightforward one because once B new arrays have been drawn from the original one, there remains the problem of choosing an *optimal* space in which to compare replicate arrays. This motivates the use of conjoint analysis as explained in section 5.

2 Confidence Regions

Construction of confidence intervals and regions is always dependent on existence of a function of the parameter of interest and the sample called the *pivot*, whose distribution is supposed to be independent of the parameter.

2.1 Percentile Method for Confidence Intervals

In order to define confidence intervals for a one-dimensional parameter $\theta = \theta(F)$, EFRON, 1981, [11] supposes that one could replace the unknown distribution function $G(x, F) = P_F(\hat{\theta} < x)$ by the bootstrap approximation $G(x, F_n)$ however recent theoretical work has shown that it is better to use bootstrap approximations to the distributions $H(x, F)$ and $K(x, F)$ of the quantities $n^{\frac{1}{2}}(\hat{\theta} - \theta)/\sigma$ and $n^{\frac{1}{2}}(\hat{\theta} - \theta)/\hat{\sigma}$ (where $n^{-1}\sigma^2$ is the asymptotic variance of $\hat{\theta}$).

Under the assumption that θ is a smooth function of the mean μ , one can show that bootstrap critical points constructed for confidence regions having a nominal probability of α are first order correct when G is used and second order correct when the studentized bootstrap is used, i.e. when built using $K(x, F)$. This encourages use of the studentized root whenever a satisfactory estimate of σ is available. However when this is not the case, iteration of the bootstrap mechanism as suggested by BERAN, 1987, [2] and HALL & MARTIN, 1988, [18], provides a way of improving poor coverage accuracy.

2.2 Confidence Regions

Generalization of the above to the multidimensional case is quite natural through use as a root of the quantity $n^{\frac{1}{2}}\hat{S}_n^{-\frac{1}{2}}(\hat{\theta} - \theta)$, however confidence regions are not uniquely defined as confidence intervals are, one possibility, as suggested by HALL, 1987, [16], is to choose the region \mathcal{C} that has minimum content subject to $P(\theta \in \mathcal{C}) = \alpha$, this is equivalent to imposing that \mathcal{C} should be such that all parameter values inside \mathcal{C} have higher likelihood than those outside. In order to do this it will be necessary to provide an estimate \hat{f} for the density

function.

Other ways of addressing this problem is to study the p -dimensional parameter as defined in polar coordinates, thus providing confidence cones for its direction.

We will see how such an approach can be useful in section 6.3.

2.3 Real Functions of a covariance-matrix

BERAN & SRIVASTAVA 1985, [5] bring down the multivariate problem to a one-dimensional one through use of the root $T_n = u(n^{\frac{1}{2}}[g(S_n) - g(\Sigma_F)])$ where

- g is a function defined from $\mathbb{R}^{p(p+1)/2} \text{ uvec}(\Sigma_F)$ in \mathbb{R}^k
- u is a real-valued function defined \mathbb{R}^k that is not constant on any set other than those having zero measure.

They show that if F has finite fourth moments and if g is continuously differentiable, then justification of the use of the bootstrap in construction of confidence regions for $g(\Sigma_F)$ is provided by the convergence of T_n^* 's distribution towards that of a transformed multi-normal distribution.

A confidence interval that has an asymptotic level of $1 - \alpha$ is provided by

$$D_{n,g}(\alpha) = \{g(\Sigma_F) : u(n^{\frac{1}{2}}[g(S_n) - g(\Sigma_F)]) \leq c_{n,g}(\alpha, \hat{F}_n)\}$$

where $c_{n,g}(\alpha, \hat{F}_n)$ is an upper α quantile of the bootstrap distribution of T_n^* .

The following examples use these results and are of practical interest :

1. $g(\Sigma_F) = \rho_{F,ij}$, confidence intervals can be provided for the correlation coefficient using the root $T_n = |r_{n,ij} - \rho_{F,ij}|$
2. In order to construct simultaneous confidence intervals for simple eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0$ BERAN & SRIVASTAVA 1985, [5] use the root

$$T_n = \max_{1 \leq i \leq p} |\log(\lambda_i(S_n)) - \log(\lambda_i(\Sigma_F))|$$

The confidence region thus obtained by resampling is

$$\{\lambda : \lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0 \text{ et } \forall i, 1 \leq i \leq p, \lambda_i(S_n)A_n^{-1} \leq \lambda_i \leq \lambda_i(S_n)A_n\}$$

with $A_n = \exp(\hat{c}_{n,g}(\alpha))$, where $\hat{c}_{n,g}(\alpha)$ is an upper α -quantile of the bootstrap distribution of the root T_n .

Justification of the use of the bootstrap in this case is provided by the fact that for any semi-definite matrix S that has simple eigenvalues the function that associates these eigenvalues to the matrix is continuously differentiable. (see KATO, [22]).

3. If F is such that the eigenvalues of Σ_F are all simple and if we denote by $\gamma_i(S)$ the i th normed eigenvector of S . Then BERAN & SRIVASTAVA use the root

$$T_n = 1 - |\langle \gamma_i(S_n), \gamma_i(\Sigma_F) \rangle|$$

in order to construct confidence cones for the eigenvectors with fixed coverage.

Remark : An interesting application of this result is the possibility of choosing in the confidence cone of a given level (for instance 95%) the vector that is the easiest to interpret

BERAN & SRIVASTAVA [5] give a good example of such an approximation.

3 Motivations and Notations for the RV-coefficient

Suppose that the data to be studied is contained in a matrix $X_{n \times p}$ of p measurements made on n observations, to which we can associate a positive semi-definite matrix $Q_{p \times p}$ that defines distances between observations :

$$d^2(x_i, x_j) = (x_i - x_j)'Q(x_i - x_j)$$

and we will also associate weights p_i to the observations these will be the elements of a diagonal matrix $D_{n \times n}$ whose trace is usually one.

Unless otherwise stated X is supposed to be centred with respect to D , i.e. $X'D1_n = 0$ where $1_n = (1, 1, \dots, 1)'$, then the variance-covariance matrix of the sample is $X'DX$.

In the subsequent sections we will always consider X together with the metrics defining distances in the observation space R^p and the variable space R^n , so we will always be analyzing triples (X, Q, D) .

Our objectives are

- 1. to simplify the triple (X, Q, D)
- 2. to compare several triples $\{(X_k, Q_k, D_k), k=1..K\}$

This will involve associating a characteristic operator to each triple, for instance when we have several arrays of different measurements made on the same observations the characteristic operator will be

$$O_k = X_k Q_k X_k' D \text{ so that } O_k \in R^{n \times n}$$

if the different studies are relative to the same variables then

$$O_k = X_k' D_k X_k Q \text{ and } O_k \in R^{p \times p}$$

4.2 PCA seen through the RV-coefficient

Find a triple $(Y_{n \times M}, I, D)$, with $M < p$ as good an approximation of rank M as possible i.e. that maximises

$$RV(O_X, O_Y) = RV(XQX'D, YY'D) = \frac{\text{Tr}(XQX'DYY'D)}{\sqrt{\text{Tr}(XQX'D)^2 \text{Tr}(YY'D)^2}}$$

This maximum is attained for Y being composed of the first M eigenvectors of $XQX'D$ such that $Y'DY = \Lambda^{(M)}$ and the value of this maximal RV-coefficient is

$$RV(O_X, O_Y) = \sqrt{\frac{\sum_{i=1}^M \lambda_i^2}{\sum_{i=1}^p \lambda_i^2}}$$

4.3 Correspondence Analysis

This is a special case of the above procedure when the data is a contingency table N and the data analysed is $X = \frac{1}{\nu}N$ where $\nu = \sum_{i=1}^n \sum_{j=1}^p n_{ij}$ is the grand total.

If r denotes the vector of row sums ($r = \frac{1}{\nu}N \cdot 1_p$), and c the vector of column sums ($c = \frac{1}{\nu}N' \cdot 1_n$), then independence is equivalent to $X = rc'$.

Correspondence analysis is equivalent to the PCA of (X, D_c^{-1}, D_r^{-1}) , where D_v^{-1} denotes the diagonal matrix whose elements are those of the inverses of the components of the vector v .

This means that the methods developed hereafter in the case of PCA can be applied in just the same way to Correspondence Analysis.

5 Analysis of a Data Cube (Conjoint Analysis)

Here we will try to simplify a situation where K different triples $\{(X_k, Q_k, D_k), k = 1..K\}$ are to be compared, where the arrays X_k must have one dimension in common, i.e. measurements on the same observations or the same variables measured on different groups of observations. The technique used here has been developed in ESCOUPIER 1987, [13], GLAÇON, 1981, [14] and LAVIT, 1988, [23]. In the following we will consider for instance that the variables remain the same throughout the K studies, i.e. $Q_k = Q$ (which is the case when the different triples are bootstrap replicates of the original one).

Then three different problems will be addressed and solved, namely representing distances between triples, summarizing the K triples and then comparing the observations and variables as seen through the different triples.

5.1 Distances between the arrays : Interstructure

Each triple is characterized by its associated operator : $O_k = X_k' D_k X_k Q$, distance between two operators can be measured through their vectoriel covariance:

$$\delta_{kl} = CovV(O_k, O_l) = Tr(O_k O_l)$$

these can be represented by a symmetric matrix $\Delta = (\delta_{kl})_{\{k=1..p, l=1..p\}}$.

Diagonalisation of Δ will provide an optimal planar representation of the arrays, as in an ordinary multidimensional scaling problem. Two equivalent triples will appear as collinear. If one wants equivalent triples to have identical projections then as in PCA one can use the reduced distances i.e. diagonalize the matrix composed of the elements

$$\rho_{kl} = RV(O_k, O_l) = \frac{\delta_{kl}}{\sqrt{\delta_{kk} \delta_{ll}}}$$

5.2 A summary of the K triples : the compromise

In principal components the objects compared are variables and the inner product is their covariance, here the objects are the operators and their inner product will be their CovV or their RV-coefficient. The first eigenvector u of the diagonalised matrix allows the construction of a linear combination

$$\tilde{O} = \sum_{k=1}^K u_k O_k$$

which has the largest norm and which maximises the sum of squares of the inner products with the initial operators.

Remark : It can be shown that all the elements of u have the same sign, so they can be chosen positive, thus \tilde{O} is a semi-definite positive operator. In fact \tilde{O} is the operator related to the triple $(\mathcal{X}, Q, \mathcal{D})$ where $\mathcal{X} = (X_1 | X_2 | \dots | X_K)$ and

$$\mathcal{D} = \begin{pmatrix} \sqrt{u_1} D_1 & & \\ & \dots & \\ & & \sqrt{u_K} D_K \end{pmatrix}$$

Diagonalization of \tilde{O} will provide a representation of the p variables, in which proximity of two points is to be interpreted as an average likeness of these variables.

5.3 Comparing the evolution of the variables and the observations through the different triples

Projection onto the first principal planes obtained by using the first two eigenvectors of \tilde{O} provides a means of comparing the variables as they are *seen* by the different triples, each variable will thus have K different projections, forming the variable's trajectory.

The same can be done for the observations, those that are not present in a triple can be represented all the same through projections, as can supplementary observations in ordinary PCA.

6 PCA and the bootstrap

6.1 Confidence Regions for the observations' representations

As is customary in Bootstrapping, from the original array X resamples X_b^* are drawn, for b varying from 1 to B , the number of bootstrap replicates needed. At the same time a new diagonal weight matrix is created through associating the weights to the observations resampled and then normalizing, so that D_b^* has unit trace.

The conjoint analysis explained above is then applied to the cube of B replicates $\{(X_b^*, Q, D_b^*), b=1..B\}$

The number M of components to be retained should already have been estimated through either cross-validation (cf HOLMES-JUNCA, 1985, [20]), or a graphical *cut-point* test, then the K different M -ranked approximations can be compared through conjoint analysis.

The representation in the first principal plane of the compromise shows for instance K different projections of each of the observations, the spread of each of these clusters will indicate the stability of the representation for each observation.

Another indicator is provided by constructing the convex hull of the projections for each observation.

Empirical confidence regions can be obtained from these clusters of B projections by constructing the successive convex hulls enveloping the points.

One could also use P. Hall's likelihood confidence region approach by doing a two dimensional density estimate for each cloud and taking the confidence region to be defined by an iso-contour for the estimated density \hat{f} . (cf. HALL 1987 [16])

This procedure provides *smoother* confidence regions without sacrificing indications of first-order departure from normality as is the case when one uses ellipses.

6.1.1 Confidence Intervals for the RV-Coefficient

The quality of a PCA can be measured, as has been already explained by the coefficient

$$RV_n^{(M)} = RV(O_n^{(M)}, O_n) = \frac{\sum_{i=1}^M \lambda_i(O_n)^2}{\sum_{i=1}^p \lambda_i(O_n)^2}$$

One could suppose that $RV_n^{(M)}$ estimates some unknown coefficient $\rho V^{(M)}$ that measures the quality of the approximation of $O^{(M)}$ to O . The most frequent case is of course when the population (unknown) covariance matrix is Σ_F and $O = \Sigma_F Q$. When all the eigenvalues of O are simple ones, $\rho V^{(M)}$ is a continuously differentiable function of the covariance matrix, and

following BERAN & SRIVASTAVA, 1985, [5] we can use the quantity $T_n = |RV_n^{(M)} - \rho V^{(M)}|$ as the root to be bootstrapped in the usual way so to obtain c_n^* such that the confidence interval :

$$D_n(\alpha) = \left\{ \rho V^{(M)} : |RV_n^{(M)} - \rho V^{(M)}| \leq c_n^* \right\}$$

has an asymptotic coverage of $(1-\alpha)$.

Remark : Other roots based on the RV coefficient can provide other bounds for the quality of a PCA's approximation are:

$$T_n = \max_{1 \leq M \leq p-1} |RV_n^{(M)} - \rho V^{(M)}|$$

Through analogy with the correlation coefficient case (HALL, MARTIN & SCHUCANY, 1989, [19]) one knows in advance that a variance stabilization transformation will give better results suggesting the roots:

$$T_n = |\operatorname{arctanh}(RV_n^{(M)}) - \operatorname{arctanh}(\rho V^{(M)})|$$

or $T_n = \left| \log \frac{1 + RV_n^{(M)}}{1 - RV_n^{(M)}} - \log \frac{1 + \rho V^{(M)}}{1 - \rho V^{(M)}} \right|$

6.2 Directional Data

6.2.1 Theory

This section details results due to DUCHARME & AL., 1985, on the construction of confidence cones for the mean direction defined as

$$\theta(F) = \frac{\mu_F}{\|\mu_F\|}$$

where μ_F denotes the mean of a random p -lengthed vector X that belongs to the p -dimensionned sphere S_p . DUCHARME & AL. construct confidence cones based on the root

$$T_n(F_n) = n(1 - \langle \theta(F), \theta(F_n) \rangle)$$

Indeed if $J_n(F_n) = \mathcal{L}(T_n(F_n))$ denotes T_n 's distribution, they show that $J_n(\hat{F}_n)$ converges to a distribution that does not depend on the direction of $\mu(F)$. In order to do this DUCHARME & AL. define a function g on \mathbb{R}^{p*} by

$$g(x) = \frac{\langle \mu(F), x \rangle}{\|\mu(F)\| \|x\|}$$

Denote by H_F g 's Hessian at μ_F , B_F the matrix of normed eigenvectors of H_F .

Then the following theorem justifies the use of the bootstrap through application of the

triangular array principle:

Theorem: Denoting by $J(F)$ the distribution of the random vector

$$\frac{1}{2\|\mu(F)\|^2} \sum_{i=2}^p Z_i^2 \text{ where } \mathcal{L}(Z_1 \dots Z_p) = \mathcal{N}(0, B'\Sigma_F B)$$

then one has the following weak convergence:

$$J_n(\hat{F}_n) \xrightarrow{p.s.} J(F)$$

This theorem implies almost sure convergence of the bootstrap critical point $c(\alpha, \hat{F}_n)$ towards the true critical point $c(\alpha, F)$, and thus validity of the resampling procedure.

The set

$$C_B = \{v \in S_p \mid \langle v, \theta(\hat{F}_n) \rangle \geq 1 - (2n)^{-1} c(\alpha, \hat{F}_n)\}$$

is a confidence region whose approximate probability coverage is $1 - \alpha$ for any cumulative distribution function F on the sphere such that $\mu(F) \neq 0$.

6.3 Confidence Cones for the Characteristic Operator

If the function g is defined as the function that associates to a definite positive symmetric matrix S , its vectoriel correlation with regards to O , $g(S) = RV(O, S)$, then g is a continuously differentiable function. Application of the preceding section's results to the root

$$T_n = n(1 - RV(O, O_n))$$

leads to confidence cones for the unknown characteristic operator O of the study without having to make any other distributional assumptions except the existence of fourth order moments.

Moreover these confidence cones also provide a way of choosing the number M of components to retain in PCA. One can choose the first integer M for which $O_n^{(M)}$ belongs to the confidence cone that has a preassigned nominal level $\alpha = 0.8$, for instance.

References

- [1] BERAN R., 1984, "*Bootstrap methods in Statistics*", Jbder. d. Dt. Math. Verein., 86, pp. 14-30.
- [2] BERAN R., 1987, "*Pre-pivoting to reduce level error of confidence sets*", Biometrika, vol.74 , pp. 457-468.
- [3] BERAN R., 1988, "*Balanced Simultaneous Confidence Sets*", J.A.S.A., vol.83, pp.679-686.
- [4] BERAN R. AND MILLAR P.W., 1986, "*Confidence Sets for a multivariate distribution*", 14, pp.431-443.
- [5] BERAN R. AND SRIVASTAVA M.S., 1985, "*Bootstrap sets and confidence regions for functions of a covariance matrix* ",Ann. Stat. 13, pp.95-115.
- [6] BUNKE O. AND DROGE B., 1984, "*Bootstrap and Cross-Validation estimates of prediction error for linear regression models*", Ann. Stat., vol. 12, pp.1400-1424.
- [7] DIACONIS P. AND EFRON B., 1983, "*Computer-Intensive Methods in Statistics*", Scientific American, vol.248, pp.116-130.
- [8] DICICCIO T. AND ROMANO J.P., 1988, "*A Review of Bootstrap Confidence Intervals*", J.R.S.S. B, 50, pp. 338-354.
- [9] DUCHARME G.R., JHUN M., ROMANO J.P., TRUONG K.N.,1985, "*Bootstrap confidence cones for directional data*", Biometrika, vol. 72, pp. 637-645.
- [10] EFRON B., 1979, "*Bootstrap Methods : Another Look at the Jackknife*", Ann. Stat., 7, pp. 1-26.
- [11] EFRON B., 1982, "*The Jackknife, the Bootstrap and Other Resampling Plans*", SIAM, Philadelphia.
- [12] EFRON B., 1987, "*Better Bootstrap Confidence Intervals*", J.A.S.A., vol. 82, pp. 171-185.
- [13] ESCOUFIER Y., 1987, "*Three-Mode Data Analysis : The statis method*",in Methods for Multidimensional Analysis, Lecture Notes from the European Courses in Statistics.
- [14] GLAÇON F.,1981, "*Analyse Conjointe de Plusieurs Matrices, Comparaison de Différentes Méthodes*", Thèse 3ème cycle, Grenoble.
- [15] HALL P.,1986, "*On the Bootstrap and Confidence Intervals*", Ann. Stat., vol. 14, pp.1431-1452.

- p.
ca,
g-
i",
or
ic-
n-
s",
ce
7,
M,
5.
or
de
4,
- [16] HALL P.,1987, "*On the Bootstrap and Likelihood-Based Confidence Regions*", *Biometrika*, vol. 74, pp.481-493.
 - [17] HALL P.,1988, "*Theoretical Comparison of Bootstrap Confidence Intervals*", *Ann. Stat.*, 16, pp. 927-953.
 - [18] HALL P. AND MARTIN M.A.,1988, "*On Bootstrap Resampling and Iteration*", *Biometrika*, 75, pp.661-671.
 - [19] HALL P.,MARTIN M.,SCHUCANY W.,1989, "*Better Nonparametric Bootstrap Confidence Intervals for the correlation coefficient*", (to appear)
 - [20] HOLMES-JUNCA S.,1985, "*Outils Informatiques pour l'Evaluation des Resultats en Analyse des Donneés*", These 3eme cycle, USTL, Montpellier.
 - [21] JOHNS V., 1988, "*Importance Sampling for bootstrap confidence intervals*", *J.A.S.A.*, vol. 83, pp.709-714.
 - [22] KATO T., 1966, "*Perturbation Theory for Linear Operators*",Springer, N.Y.
 - [23] LAVIT C., 1988, "*Analyse Conjointe de Tableaux Quantitatifs*", Ed. Masson.
 - [24] ROBERT P. AND ESCOUFIER Y., 1976, "*A Unifying Tool for Linear Multivariate Methods: the RV-coefficient*", *Applied Statistics*, vol.25, pp. 257-265.